Hassan Ghaffari

+989358884421 ● hasanghaffari93@gmail.com ● linkedin.com/in/ghaffarisaravi ● github.com/hasanghaffari93

Artificial Intelligence Engineer

AI engineer with +6 years of experience building and deploying real-time computer vision and NLP systems in startup environments. Hands-on with LangChain, RAG, and LLM-based chatbots; recently developed agentic Q&A apps and integrated AI into games and educational tools. Proficient in PyTorch, TensorFlow, and on-device model optimization.

WORK EXPERIENCE

Self-Employed – Tehran, Iran April 2025 - Current

AI Engineer

Developed an AI-driven chatbot providing accessible and personalized support for psychiatric-related queries

- Developed LLM-powered chatbots using LangChain with FastAPI backend for scalable API services.
- Designed stateful agentic RAG using LangGraph with Pinecone vector search, tool-calling
- Supported multiple LLM providers, input/output audio (OpenAI Whisper/TTS), and MongoDB for persistence
- Implemented document summarization using both Stuff and Map-Reduce techniques.

Phiji Play Inc – Tehran, Iran January 2022 - April 2025

Lead AI Engineer

Responsible for AI in an educational hybrid physical-digital gaming system for children (Link).

- Directed training, evaluation, and deployment of CV models for object detection and face recognition.
- Managed Unity, Android, IOS, and software teams to integrate CV/NLP models into production.
- Developed a context-aware Q&A system using RAG to enhance children's learning in interactive storytelling.
- Led data pipeline for collecting and annotating 3M+ high-quality real-world images.
- Optimized CV models for real-time offline deployment on Android and IOS devices (10-20 FPS).
- Improved efficiency of and deployed real-time offline speech recognition model (Kaldi).

Robokids – Tehran, Iran May 2021 - October 2023

AI Engineer

Developed AI solutions for interactive smart games at robotic amusement parks, Robokids (Link).

- Trained, optimized, and deployed child re-identification (ReID) models in real-time (60+ FPS).
- Conducted data collection/annotation for ReID in a real-world setting (10k+ images of 400 IDs).
- Collaborated with Unity and Backend teams to integrate AI into games.
- Implemented image processing techniques, real-time pose estimation and segmentation for 6 smart games.
- Developed a custom face image generation system using Stable Diffusion 1.5 and Lora fine-tune method, enabling personalized AI-generated portraits from just 3-6 reference photos.

Institute for Cognitive Science Studies – Tehran, Iran September 2019 - September 2021

Graduate Research Assistant

- Analyzed EEG data from 100+ participants to investigate gamma-band ASSR as a potential biomarker for prediction of clinical response to rTMS in treatment resistant depression (TRD).
- Applied advanced signal processing techniques (ERSP, ITC) and statistical analysis to uncover neural-cognitive associations in healthy and clinical populations.

Jarfabin – Tehran, Iran May 2019 - May 2021

Computer Vision Researcher and Developer

- Designed, developed, and implemented multi-view deep CNN models to classify pistachio nut.
- Managed pipelines of data collection, preprocessing, and annotation of 4-view pistachio nuts (1M+).
- Optimized CV models for real-time inference on edge (Raspberry Pi and Jetson Nano).

SKILLS

Programming Languages: Python (expert) | Matlab (expert) | Rust/C++ (basic proficiency)
ML Frameworks: TensorFlow | PyTorch | ONNX | TensorRT | OpenVino
LLMs: LangChain | LangGraph | AI Agents | RAG | Chatbots | OpenAI API | Groq
Tools: OpenCV | Socket | Git | Docker | Scikit-learn | Matplotlib
Databases: noSQL (MongoDB) | Vector Stores (Pinecone)
Vision Tasks: Model Optimization | Object Detection | Segmentation | Re-identification

EDUCATION

Institute for Cognitive Science Studies – Tehran, Iran
Master of Arts - Cognitive Science September 2021

- GPA: 19.33 out of 20 (4/4)
- Advisor: Dr. Fidel Vila-Rodriguez (UBC)
- Thesis: Auditory steady-state response (ASSR) as predictor of clinical response to rTMS in treatment-resistant depression (TRD) and cognitive performance

Ferdowsi University of Mashhad – Mashhad, Iran
Bachelor of Science - Electrical Engineering September 2016

- GPA: 15.49 out of 20

CERTIFICATES

Machine Learning – Stanford University November 2020
Deep Learning Specialization – Coursera September 2020
TensorFlow Developer Specialization – Coursera October 2020
AI for Medicine Specialization – Coursera November 2020

LANGUAGE

Persian: native language
English: C1 (IETLS 7.0, listening 7, reading 8.5, writing 6.5, speaking 6, June 2024)